Data centers in space, AI's energy needs and musical phones

IBM Think Newsletter <IBM@email.ibm.com>
Reply-To: ibmcorpmain@ibm.com
To: ajdellinger@gmail.com

Tue, Nov 19, 2024 at 8:23 AM

View this email within a webpage

Think Newsletter

IBM

AI needs energy. But maybe not as much as you think.

A note from this week's editor, reporter AJ Dellinger:

As AI adoption grows, so do the energy needs of data centers. Goldman Sachs Research predicts that power demand from this sector will increase by 160% by 2030, which means solutions are needed—and fast.

While AI has been dinged from some corners for large energy consumption, it doesn't have to be that way. One option, experts say, is AI itself, which in theory could help optimize energy efficiency, lower data processing costs and assist in clean energy production.

Phil Spring, EMEA Energy & Resources Industry Leader at IBM Consulting, offers several solutions for keeping AI development sustainable:

— **Test foundation models**: "[They] can be fine-tuned for specific

purposes, but only require training once."
— **Consider smaller models**: "Smaller models trained on high-quality, curated data can often achieve similar or even better results [than LLMs], while being more energy-efficient."
— **Choose your processing location carefully**: "Situating data centers in areas with renewable energy sources can lead to significant energy savings."

Get the story →

AI and model news

At what point will AI scale break?

Bigger doesn't always mean better. And for AI models, there comes a tipping point where training on more data doesn't yield better results. Are we there yet? Find out on *Mixture of Experts*.

See the episode →

In five years, AI has seen some stuff. Here's what IBM's AI Ethics board makes of it.

The future of AI is literally wide open, and the way forward is rooted in responsibility. Explore how AI governance, regulation, transparency and

trust have evolved over the past five years—and what this means for your business.

Get the insights →

AI for small business: Your new co-founder

Whether it's serving as a sounding board, a content creation engine or your personal research assistant, gen AI is redefining entrepreneurship—helping small business owners go to market and profit faster.

Discover how →

More AI news

IBM announced that its most advanced quantum processor, the **Quantum Heron**, can now run certain quantum circuits with up to 5,000 two-qubit operations and execute experiments up to 50 times faster than previous models.

Researchers unveiled **FrontierMath**, a new **advanced mathematical reasoning benchmark** they say exposes a significant gap between current AI capabilities and human mathematical expertise.

IBM has become **UFC**'s first-ever Official Global AI Partner, with plans to develop a **UFC Insights Engine** that uses **IBM watsonx** to provide fans with real-time fight analysis and predictions.

IBM Research's Docling, an open source project and tool for parsing documents, is **GitHub**'s top repository this week.

Agmatix's AI assistant, **Leafy**, uses **Amazon Bedrock** to help farmers

analyze field trial data faster, accelerating the development of better seeds and crop protection products.

Anthropic reportedly hired its first **AI welfare researcher** to develop frameworks to assess whether AI is capable of consciousness and suffering, aiming to guide ethical decisions at Anthropic and beyond.

Skill up: **IBM SkillsBuild** offers **free AI courses** and credentials to help workers adapt to a job market where 92% of ICT positions require new technology skills.

Security news

Balance due: Your login credentials

New research from IBM X-Force shows an uptick in sophisticated phishing campaigns using real, stolen invoice notifications to appear more legitimate. Here's what you need to know:

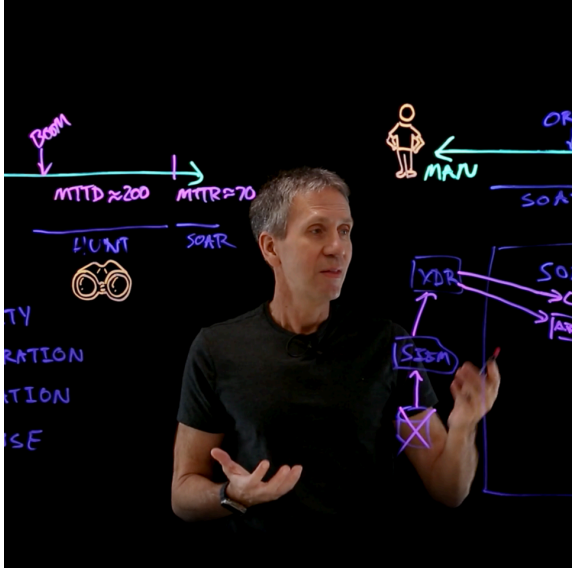
— Hive0145, a financially motivated initial access broker, is believed to be the group behind the ongoing attacks.
— The group deploys Strela Stealer malware, which is designed to extract users' valid email credentials.
— Several commercial organizations in Europe have been targeted across industries, including finance, e-commerce and manufacturing.
— Hive0145's use of techniques like email hijacking and obfuscated file attachments indicates a shift in its maturing cyber operations capability.

Read on for the full analysis.

Get the insights →

Meet the 25-year-old security threat that refuses to die

Cross-site scripting, or XSS, has been around since 1999 but remains one of the most impactful security attack types. How can you minimize this threat? Validate site inputs and encode outputs.



Jeff Crume explains →

More security news

Microsoft published 89 security issues in its **November Patch Tuesday** release, including two zero-days with active exploitations detected and a high-severity Exchange Server vulnerability that makes it easier to create effective malicious emails.

A **MOVEIt data leak** released a large amount of **sensitive employee data** from major organizations like Amazon, HSBC, British Telecom and McDonald's to the dark web, including contact information, job titles and internal structures.

The **US Department of Justice** indicted and charged two individuals believed to be the hackers behind the **Snowflake data breach**, which resulted in the theft of 50 billion AT&T customer call and text records and the alleged extortion of about USD 2.5 million from three victims.

Sustainability news

IBM and Sustainable Energy for All team up on two new AI models for urban development

IBM and **Sustainable Energy for All** have jointly launched two models that aim to help policymakers identify energy and infrastructure needs for cities and communities in developing regions.

Meet the models:

— **Open Building Insights**: A platform that helps policymakers and other stakeholders consolidate data about building usage in an area so they can make decisions that lead to more sustainable development, such as determining which communities need electrification and energy services.
— **Modeling Urban Growth**: An open-source AI model available on GitHub that can predict where cities will grow—and help decision-makers and developers plan for future urbanization needs.

Get the news →

More sustainability news

The **US**, along with several other countries, announced a new framework at **COP29** that would seek to **triple each nation's nuclear energy capacity** by 2050. The goal for participating countries would be to introduce 15 gigawatts of nuclear power by 2035, and an additional 15 gigawatts per year by 2040.

Also at COP29, the **UK** unveiled a framework for **establishing a voluntary market for nature credits**, which could enable governments and private sector players to produce and trade securities that each represent a unit of improvement to biodiversity.

New research from **McKinsey Global Institute** suggests that world governments are **lagging behind in clean energy adoption**, deploying just 10% of the low-emissions technologies needed to reach 2050 carbon goals.

Cloud news

Want storage in the cloud?
Aim higher.

The data centers that fuel AI applications consume an enormous amount of electricity on Earth. So why not harness solar energy directly by launching them into orbit?

That's the elevator pitch of Lumen Orbit, a Y Combinator-backed startup that has developed a moonshot plan to build a network of data centers in space. Lumen is gearing up to launch demonstration satellite technology as early as 2025, with support from the NVIDIA Inception startup incubator.

Though it sounds like sci-fi, researchers have been investigating the possibility of orbital data centers for a while now. And while approaches differ, one thing is clear: there will soon be a cloud above the clouds.

Get the story →

More cloud news

Red Hat announced that it will acquire **Neural Magic**, a leader in AI inference acceleration software, to strengthen its high-performance AI capabilities across hybrid cloud environments.

IBM is showcasing its **high-performance computing (HPC)** solutions at **SC24** in Atlanta through November 22, demonstrating how advanced interconnects and accelerators power AI and simulation workloads.

LocalStack has secured USD 25 million in a funding round led by **Notable Capital**. The Swiss startup offers an open-source tool that lets developers simulate an AWS environment on their local machines.

Google announced that **Google Kubernetes Engine (GKE)** is ready to handle trillion-parameter AI models.

Final thoughts

Everybody knows you can play music on your smartphone. But can you turn your smartphone into an instrument? Let's examine the evidence.

Zefiro is a tiny **MIDI wind controller** that you plug into your USB-C port and blow into. It sounds like a recorder, kind of. It definitely looks like a vape.

A spiritual precursor to Zefiro was an app called **Ocarina**, which "literally turns your **iPhone into a 12,000-year-old, flute-like wind instrument**" according to its creator. (You blow into the port directly, which seems like it'd void your AppleCare.)

Finally, the short-lived **Artiphon Instrument 1** was a rectangular **MIDI controller** that lets you "strum, slide, tap and drum any sound." It looked like a little Chapman Stick, but sadly did not make you sound like Tony Levin from King Crimson.

Was this email forwarded to you? You can subscribe here

© Copyright International Business Machines Corporation 2024.

IBM

You are receiving this email because you are signed up for: IBM Think Newsletter

Stop receiving the IBM Think Newsletter

Manage subscription

Unsubscribe from other IBM marketing communications

IBM's Privacy Statement

Trademarks

IBM US, 1 North Castle Drive, Armonk, NY 10504, United States